LIMIT THEOREMS FOR FUNCTIONS OF SHORTEST TWO-SAMPLE SPACINGS AND A RELATED TEST[1]

By Saul Blumenthal

New York University

1. Introduction and notation. Limit theorems for certain functions of two-sample "sample spacings" are given, and then applied to obtain the large sample properties of a procedure for testing whether two distribution functions ($F(x)$ and $G(x)$) are the same. The present limit results extend earlier work of Blum and Weiss [1], and the proposed test is analogous to one used by Weiss [6].

Denote observations from one population by $X_1, X_2, \cdots, X_m$ and from the other population by $Y_1, Y_2, \cdots, Y_n$, with labels chosen so that $m = \theta n$ with $\theta \geqq 1$. The X's are independent with common distribution function $F(x)$, and the Y's are independent with common distribution function $G(x)$. Let p_0 ($0 < p_0 < 1$) be given (choice of a value for p_0 will be discussed in Section 3).

The ordered X-values will be denoted $X_1' \leqq \cdots \leqq X_m'$, and the ordered Y's by $Y_1' \leqq \cdots \leqq Y_n'$. Let Y_0' denote $-\infty$ and Y'_{n+1} denote $+\infty$. By S_i we denote the number of $X_1, \cdots, X_m$ which are contained in the interval $[Y'_{i-1}, Y_i')$ ($i = 1, \cdots, n+1$). The S_i are the numbers of X's "separating" adjacent ordered Y's and are sometimes referred to as "sample spacings." S_i will be seen to be a measure of the "probability content" of the interval $[Y'_{i-1}, Y_i')$.

For an arbitrary k and collection of indices $(i_1, \cdots, i_k)$ we write

$$I_n = \textstyle\bigcup_{j=1}^{k} [Y'_{i_j-1}, Y'_{i_j}) \tag{1.1}$$

and we denote the "content" of I_n as

$$H_n = \textstyle\sum_{j=1}^{k} (S_{i_j} + 1)/(n + m + 1). \tag{1.2}$$

We shall study $I_n(p_0)$ where the indices i_j are chosen so that intervals $[Y'_{i-1}, Y_i')$ with small corresponding S_i values are included in $I_n(p_0)$, and enough intervals are included so that $H_n(p_0)$ is as close to p_0 as possible without exceeding p_0. Thus if any interval with an S_i value of r is included in $I_n(p_0)$, all intervals with S_i values of less than r will be included. Generally many intervals will have a given S_i value, and if inclusion of all intervals with $S_i = r_0$ (say) would make $H_n(p_0) > p_0$, then an arbitrary subset of those intervals can be chosen subject to

$$p_0 - [(r_0 + 1)/(n + m + 1)] < H_n(p_0) \leqq p_0. \tag{1.3}$$

To formalize the definition of $I_n(p_0)$, we define K_n as the largest integer such

Received 18 June 1965; revised 11 August 1966.

[1] Part of this work is contained (slightly modified) in a thesis submitted to Cornell University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. Research supported in part by Air Force Contract No. AF49(683)-230, Cornell University, and by National Science Foundation Grant GP4933, Rutgers—The State University.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.*
www.jstor.org

that

$$(1.4)\quad \sum_{[i:S_i \leqq K_n]} ((S_i + 1)/(n + m + 1)) \leqq p_0 < \sum_{[i:S_i \leqq K_n+1]} ((S_i + 1)/(n + m + 1)).$$

Further, define L_n by

$$(1.5)\quad L_n(K_n + 2) \leqq (n + m + 1)p_0 - \sum_{[i:S_i \leqq K_n]} (S_i + 1) < (L_n + 1)(K_n + 2).$$

Then $I_n(p_0)$ is the union of all intervals $[Y'_{i-1}, Y'_i)$ with S_i values $\leqq K_n$ and L_n of the intervals with S_i values of $K_n + 1$, chosen at random or by a convention such as including the L_n associated with the smallest values of Y'_i.

Let E_n be the event which occurs if and only if an arbitrarily chosen Y-value (say Y_{n+1}) falls in the region $I_n(p_0)$.

In the next section we shall show that the probability of E_n converges wp 1 as n increases to $\bar{P}$ and that $\bar{P} \geqq p_0$, with equality if and only if $F(x) = G(x)$ (a.e. x). This fact will be used to construct a two-sample test in Section 3.

2. Convergence results. The quantity of interest is

$$(2.1)\qquad p_n = P(E_n) = \sum_{[i:[Y_{i-1}, Y_i')\varepsilon I_n]} (G(Y_i') - G(Y'_{i-1})).$$

We shall show that p_n converges wp 1 to a specified constant $\bar{P}$ (given by (2.27)).

The method of attack will be to study the quantities

$$(2.2)\qquad p_n(r) = \sum_{[i:S_i=r]} (G(Y_i') - G(Y'_{i-1})),$$

i.e. the probability assigned to the Y-spacings containing exactly r X's. Note that

$$(2.3)\qquad p_n = \sum_{r=0}^{K_n} p_n(r) + p_n(K_n + 1, L_n)$$

where

$$(2.4)\quad p_n(K_n + 1, L_n) = \sum_{j=1, [i_j:S_{i_j}=K_n+1]}^{L_n} (G(Y'_{i_j}) - G(Y'_{i_j-1})).$$

It will be necessary in establishing the convergence of p_n to show that the four quantities $p_n(r)$, K_n, L_n, and $p_n(r, L)$ all converge (K_n and L_n are defined by (1.4) and (1.5) respectively).

Before undertaking that task, we note that

$$(2.5)\qquad G(Y_i') = GF^{-1}(F(Y_i'))$$

assuming that $F^{-1}(x)$ is well defined. Note also that the quantities S_i are unchanged under a monotone transformation such as that which sends all X_i into $F(X_i)$ and all Y_i into $F(Y_i)$. The quantities $F(X_i)$ have the uniform distribution $U(x)$ on $(0, 1)$. Thus p_n is the same if $F(x)$ and $G(x)$ are the distributions of the X's and Y's respectively as if $U(x)$ and $GF^{-1}(x)$ are the distributions of the X's and Y's respectively. For notational convenience, *throughout the remainder* *of this section, we assume the X's to have the uniform distribution and write the distribution $GF^{-1}(x)$ of the Y's simply as $G(x)$.*

The first step will be to establish the stochastic convergence of K_n and L_n. Define

$$(2.6)\quad Q_n(r) = (1/n+1)\,(\text{the numbers of } S_1, \cdots, S_{n+1} \text{ which equal } r),$$

(i.e., the proportion of the Y-spacings which contain exactly r values of X.) The stochastic convergence of $Q_n(r)$ has been established by Blum and Weiss [1] and will be stated here.

THEOREM 2.1. *Define*

$$(2.7)\qquad Q(r) = \theta^r \int_0^1 g^2(y)(\theta + g(y))^{-(r+1)}\, dy.$$

Given $\epsilon, \delta > 0$ and $R > 0$ (R an integer), there exists $N(\epsilon, \delta)$ such that

$$(2.8)\qquad P[\sup_{r \geqq 0}|Q_n(r) - Q(r)| < \epsilon, \quad \text{all} \quad n > N(\epsilon,\delta)] \geqq 1 - \delta,$$

$$(2.9)\quad P[|\textstyle\sum_{r=0}^R rQ_n(r) - \sum_{r=0}^R rQ(r)| < \epsilon, \quad \text{all} \quad n > N(\epsilon,\delta)] \geqq 1 - \delta$$

where $N(\epsilon, \delta)$ does not depend on R.

PROOF. Statement (2.8) is proved by Blum and Weiss [1], and a slightly modified version of (2.8) is derived by this author in [2]. Statement (2.9) follows from (2.8) and

$$(2.10)\quad (n + 1/n) \textstyle\sum_{r=0}^{\infty} rQ_n(r) = (1/n) \sum_{i=1}^{n+1} S_i = \theta = \sum_{r=0}^{\infty} rQ(r),$$

completing the proof.

Next, we define a quantity analogous to the "sample content" H_n (see (1.2)), namely

$$(2.11)\quad H_n(K) = \textstyle\sum_{[i:S_i \leqq K]} (S_i + 1)/(n(\theta + 1) + 1)$$
$$= \textstyle\sum_{r=0}^K (r + 1)Q_n(r)((n + 1)/(n(\theta + 1) + 1))$$

where K is a fixed integer. Further, define

$$(2.12)\qquad H(K) = \textstyle\sum_{r=0}^K ((r + 1)/(\theta + 1))Q(r).$$

An immediate corollary to Theorem 2.1 is the convergence of $H_n(K)$ to $H(K)$.

COROLLARY 2.1 *Given $\epsilon, \delta > 0$, and a positive integer K,*

$$(2.13)\qquad P[|H_n(K) - H(K)| < \epsilon, \quad \text{all} \quad n > N(\epsilon,\delta)] \geqq 1 - \delta.$$

This corrollary will be used to demonstrate the stochastic convergence of K_n, as follows. Define the integer K_0 by

$$(2.14)\qquad H(K_0) \leqq p_0 < H(K_0 + 1),$$

which is a direct analogue of the defining equation (1.4) of K_n.

COROLLARY 2.2 *With K_n and K_0 defined by (1.4) and (2.14) respectively, and if $H(K_0) < p_0$, and given δ,*

$$(2.15)\qquad P[K_n = K_0 \quad \text{all} \quad n > N(\delta, p_0, G)] \geqq 1 - \delta,$$

where $N(\delta, p_0, G)$ *depends on* δ, *the parameter* p_0 *and the distribution* G *but on nothing else. If* $H(K_0) = p_0$, *then*

$$P[K_n = K_0 \text{ or } K_0 - 1, \quad \textit{all} \quad n > N(\delta, p_0, G)] \geqq 1 - \delta. \tag{2.16}$$

PROOF. For (2.15), take $\epsilon = \min((p_0 - H(K_0))/2, (H(K_0 + 1) - p_0)/4)$ in (2.13). Since K_0, $H(K_0)$ and p_0 depend on p_0 and G (only), (2.15) follows. For (2.16), take $\epsilon = \min((p_0 - H(K_0 - 1))/4, (H(K_0 + 1) - p_0)/4)$. This completes the proof.

The case $H(K_0) = p_0$ introduces slight complications into the argument which can only obscure the general ideas, so only the case $H(K_0) < p_0$ will be considered. The modifications for the other case should become apparent.

Next, define L_0 (an integer) as

$$L_0 = (p_0 - H(K_0))(\theta + 1)/(K_0 + 2). \tag{2.17}$$

COROLLARY 2.3. *With* L_n *defined by* (1.5) *and* L_0 *by* (2.17), *given* $\epsilon, \delta > 0$,

$$P[|(L_n/n) - L_0| < \epsilon, \quad \textit{all} \quad n > N(\epsilon, \delta, p_0, G)] \geqq 1 - \delta. \tag{2.18}$$

The proof follows easily from the definition (1.5) and the previous two corollaries.

With the convergence of K_n and L_n established, the convergence of $p_n(r)$ (see (2.2)) will be demonstrated next.

THEOREM 2.2. *Define*

$$P(r) = (r + 1)\theta^r \int_0^1 (g^3(y)/(\theta + g(y))^{(r+2)})\, dy. \tag{2.19}$$

Given $\epsilon, \delta > 0$ *and* R *a positive integer,*

$$P[\sup_{r \geqq 0}|p_n(r) - P(r)| < \epsilon, \quad \textit{all} \quad n > N(\epsilon, \delta)] \geqq 1 - \delta, \tag{2.20}$$

$$P[|\textstyle\sum_{r=0}^{R} p_n(r) - \sum_{r=0}^{R} P(r)| < \epsilon, \quad \textit{all} \quad n > N(\epsilon, \delta)] \geqq 1 - \delta, \tag{2.21}$$

where the constant $N(\epsilon, \delta)$ *depends only on* ϵ *and* δ, *not on* R.

PROOF. The proof of (2.20) is contained in [2] and will not be given here. It can be obtained using (2.8) and the main result of Weiss [5].

Expression (2.21) is derived easily from (2.20) and

$$\textstyle\sum_{r=0}^{\infty} p_n(r) = 1 = \sum_{r=0}^{\infty} P(r). \tag{2.22}$$

Next, consider for $0 < \lambda < 1$,

$$p_n(r, \lambda(n + 1)Q_n(r)) = \textstyle\sum_{j=1,[i_j:S_{i_j}=r]}^{\lambda(n+1)Q_n(r)} (G(Y'_{i_j}) - G(Y'_{i_j-1})). \tag{2.23}$$

Note the similarity between (2.23) and (2.4).

COROLLARY 2.4. *Given* $\epsilon, \delta > 0$,

$$P[\sup_{r \geqq 0}|p_n(r, \lambda(n + 1)Q_n(r)) - \lambda P(r)| < \epsilon, \quad \textit{all} \quad n > N(\epsilon, \delta)] \geqq 1 - \delta. \tag{2.24}$$

PROOF. From the proof of (2.20), it can be seen that (2.20) remains true for $n' = \lambda n$ with the obvious insertions of λ's as in (2.24).

Finally, the convergence of p_n can be demonstrated.

THEOREM 2.3. *Given* $\epsilon, \delta > 0$,

$$(2.25)\quad P[|p_n - \sum_{r=0}^{K_0} P(r) - (L_0/Q(K_0 + 1))P(K_0 + 1)| < \epsilon, \quad all \quad n > N(\epsilon, \delta, p_0, G)] \geqq 1 - \delta.$$

PROOF. Using (2.3) to represent p_n, the theorem follows from the straightforward application of Theorems 2.1 and 2.2 and Corollaries 2.2, 2.3, and 2.4.

COROLLARY 2.5. *Given* $\epsilon > 0$,

$$(2.26)\quad P[|p_n - \sum_{r=0}^{K_0} P(r) - (L_0/Q(K_0+1))P(K_0+1)| > \epsilon] < C(\epsilon)/n^2.$$

PROOF. It is not difficult to establish that the $2r$th moment of $(p_n - \sum_{r=0}^{K_0} P(r) - (L_0/Q(K_0 + 1))P(K_0 + 1))$ is $O((n^{-\frac{1}{2}})^{2r})$. This is seen during the proof of Theorem 2.3 where the case $r = 1$ is carried out in detail. (See [2]). To get (2.26), take $r = 2$ and use the generalized Chebychev inequality.

Stronger results than Corollary 2.5 are obtainable (e.g. n^2 could be replaced by n^r, any $r > 0$) but are not needed here. Regarding the behavior of p_n, it is seen from (2.1) that p_n is a sum of a random subset of a set of interchangeable random variables (namely the $\{G(Y_i') - G(Y_{i-1}')\}$). Except for the indices in the summation being chosen at random, the results of Hanson and Koopmans [4] on exponential convergence rates for sums of interchangeable variables would apply here and we believe that these rates do apply to these p_n. Also, the central limit theorem on Chernoff and Teicher [3] for sums of interchangeable variables "almost" applies here (with the same exception) and again we believe that these p_n do obey the central limit theorem. These properties will not be studied in this paper.

We shall show now that p_n tends to p_0 when $F(x) = G(x)$ and to a greater value otherwise. Let $\bar{P}$ be defined by

$$(2.27)\quad \bar{P} = \sum_{r=0}^{K_0} P(r) + (L_0/Q(K_0 + 1))P(K_0 + 1).$$

Theorem 2.3 demonstrated the almost sure convergence of p_n to $\bar{P}$. It will be shown now that $\bar{P} \geqq p_0$ with equality if and only if $F(x) = G(x)$ (e.g. $G(x)$ is uniform).

LEMMA 2.1. *Let* $\varphi(t)$ *be a positive, strictly decreasing function of* $t(t \geqq 0)$, *and let* $g(x)$ *be a continuous density function on* [0,1].

$$(2.28)\quad \int_0^1 \varphi(g(x))(1 - g(x))\, dx \geqq 0$$

with equality if and only if $g(x) = 1$ (a.e.) *on* [0, 1].

PROOF. Write $[0, 1] = S_0 U S_1$ where the mutually exclusive sets S_0 and S_1 are

$$(2.29)\quad S_0 = [x: g(x) < 1];$$

$$S_1 = [x: g(x) \geqq 1].$$

Clearly, since $\varphi(t)$ is decreasing

$$(2.30)\quad \int_{S_0} \varphi(g(x))(1 - g(x))\, dx \geqq \varphi(1) \int_{S_0} (1 - g(x))\, dx$$

with equality if and only if S_0 has measure zero. Similarly, since $(1 - g(x))$ is negative on S_1 and $\varphi(t)$ decreases,

$$\int_{S_1} \varphi(g(x))(1 - g(x))\, dx \geqq \varphi(1) \int_{S_1} (1 - g(x)\, dx \tag{2.31}$$

with equality if and only if the set $[x:g(x) > 1]$ has measure zero. Adding (2.30) and (2.31) gives (2.28). This proves the lemma.

THEOREM 2.4. *With $\bar{P}$ defined by* (2.27),

$$\bar{P} \geqq p_0 \tag{2.32}$$

with equality if and only if $G(x)$ is the uniform distribution on [0, 1].

PROOF. Using the definitions (2.12), (2.14), and (2.17) of $H(K)$, K_0, and L_0 respectively, p_0 can be expressed as

$$p_0 = (L_0/Q(K_0 + 1))H(K_0 + 1) + (1 - (L_0/Q(K_0 + 1)))H(K_0). \tag{2.33}$$

Rewrite (2.27) as

$$\bar{P} = (L_0/Q(K_0 + 1)) \textstyle\sum_{r=0}^{K_0+1} P(r) + (1 - (L_0/Q(K_0 + 1))) \sum_{r=0}^{K_0} P(r). \tag{2.34}$$

Comparing (2.34) and (2.33), (2.32) follows from the fact that

$$H(K) \leqq \textstyle\sum_{r=0}^{K} P(r) \tag{2.35}$$

for any K with equality if and only if $G(x)$ is the uniform distribution on [0,1]. To establish (2.35), use (2.7), and (2.12) to show

$$H(K) = 1 - (\theta^{K+1}/(1 + \theta)) \int_0^1 (\theta + (K + 2)g(x))/(\theta + g(x))^{K+1}\, dx. \tag{2.36}$$

From (2.19), comes

$$\textstyle\sum_{r=0}^{K} P(r) = 1 - \theta^{K+1} \int_0^1 (\theta + (K + 2)g(x))g(x)/(\theta + g(x))^{K+2}\, dx. \tag{2.37}$$

Using (2.36) and (2.37), (2.35) becomes

$$(\theta^{K+2}/(1 + \theta)) \int_0^1 ((\theta + (K + 2)g(x))/ (\theta + g(x))^{K+2})(1 - g(x))\, dx \geqq 0. \tag{2.38}$$

But (2.38) follows from Lemma 2.1 with $\varphi(t) = (a + bt)/(a + t)^b$ $(a, b < 1)$. This completes the proof.

3. A two sample test. We shall now use the previous convergence results to construct a statistic for testing the hypothesis

$$H_0:F(x) = G(x) \tag{3.1}$$

against general alternatives.

Assume for convenience that θ is an integer, and that $X_1, \cdots, X_{n\theta}$; and $Y_1, \cdots, Y_{n+1}$ have been observed. Let p_0 be fixed. On the basis of $X_1, X_2, \cdots, X_\theta$ and Y_1, form the region $I_1(p_0)$. Define

$$\delta_1 = 1 \quad \text{if } Y_2 \text{ is in } I_1(p_0) \tag{3.2}$$
$$= 0 \quad \text{otherwise.}$$

Then form $I_2(p_0)$ based on $(X_1, \cdots, X_{2\theta})$, (Y_1, Y_2) and define δ_2 in terms of Y_3 and $I_2(p_0)$. Successively, $\delta_3, \cdots, \delta_n$ are defined in this manner. Let

$$D_n = \sum_{i=1}^{n} \delta_i . \tag{3.3}$$

Thus D_n is the number of occurrences of the event E_i (see Section 1) in n trials—which are not independent.

Using Theorem 2.3, it is easily verified that $E(D_n/n)$ converges to $\bar{P}$ (given by (2.27)) as n increases. Further, it can be shown that $E((D_n/n) - \bar{P})^2$ converges to zero as n increases. The demonstration is a direct parallel of that in Section 3 of Weiss [6] and will be omitted. Thus (D_n/n) converges stochastically to $\bar{P}$. In view of Theorem 2.4, it is seen that a test which rejects H_0 for large values of $[(D_n/n) - p_0]$, and accepts otherwise, will be consistent.

To find the approximate critical values for (D_n/n) the limiting distribution of D_n is needed. It is conjectured that

$$n^{\frac{1}{2}}[(D_n/n) - \bar{P}]/(\bar{P}(1 - \bar{P}))^{\frac{1}{2}}$$

is approximately normally distributed for large n.

The normality assumption is made plausible by noting that for large values of i, $\delta_i, \delta_{i+1}, \cdots, \delta_n$, is a sequence of Bernoulli random variables, which though not independent all have (approximately) the same probability $\bar{P}$ of attaining the value unity. The fact that the limiting variance does not reflect the dependence may be justified by the following heuristics: The covariance of δ_i, δ_{i+1} depends on terms such as $P(\delta_i = 1, \delta_{i+1} = 1)$, and thus on $P(\delta_{i+1} = 1 \mid \delta_i = 1)$. The condition affects the probability involved because $\delta_i = 1$ means that the S_i value of one interval contained in $I_i(p_0)$ is increased, which could mean that in forming $I_{i+1}(p_0)$ one interval contained in $I_i(p_0)$ might be forced out and replaced by another interval. If two intervals are interchanged because of this condition, p_{i+1} would change from its unconditioned value by the difference between the probability contents of the intervals involved. Since these contents are of order $(1/i)$, the difference is of order $(1/i)^2$, and is thus relatively negligible.

On the normality assumption, the critical region of size α will be approximately

$$(D_n/n) > p_0 + (K_\alpha(p_0(1 - p_0))^{\frac{1}{2}}/n^{\frac{1}{2}}), \tag{3.4}$$

where $\Phi(K_\alpha) = 1 - \alpha$ and $\Phi(\cdot)$ is the standard normal cdf. Further, the approximate power of this test will be

$$1 - \Phi\{n^{\frac{1}{2}}[(p_0 - \bar{P})/(\bar{P}(1 - \bar{P}))^{\frac{1}{2}}\} + K_\alpha((p_0(1 - p_0))^{\frac{1}{2}}/(\bar{P}(1 - \bar{P}))^{\frac{1}{2}})\}. \tag{3.5}$$

To consider rational choices of p_0 and θ, we shall make an approximate evaluation of the power when $F(x)$ is the uniform distribution on [0, 1] and $G(x)$ is "close" to $F(x)$ and has density $g(x)$ given by

$$g(x) = 1 + ch(x), \qquad 0 \leqq x \leqq 1, \tag{3.6}$$

where,

(3.7) $ch(x) > -1; \int_0^1 h(x)\,dx = 0; \int_0^1 h^2(x)\,dx = D < \infty.$

Also assume

(3.8) $\lim_{c\to 0} C^{K-2} \int_0^1 h^K(x)\,dx = 0$ uniformly in $K(>2)$.

This assumption makes approximate computing formuli easier to obtain.

Use (3.6) for $g(x)$ in the formula (2.36) for $H(K)$, and simplify by means of (3.8) to obtain

(3.9) $H(K) = 1 - \theta^{K+1}(\theta + K + 2)/(1+\theta)^{K+2} + O(c^2), \qquad (c \to 0).$

Thus $H(K)$ differs under $G(x)$ by $O(c^2)$ from its value under $F(x)$. With K_0 defined by (2.14), and K^* defined by

(3.10) $$\theta^{K^*+2}(\theta + K^* + 3)/(1+\theta)^{K^*+3} < 1 - p_0 \leqq \theta^{K^*+1}(\theta + K^* + 2)/(1+\theta)^{K^*+2},$$

(3.9) implies that $K_0 = K^*$ for $c < c^*$, where c^* depends on p_0, $h(x)$ and θ.

With $\bar{P}$ defined by (2.27), using the definitions (2.7), (2.12), (2.14), (2.17), and (2.19) gives

(3.11) $$\bar{P} = p_0 + (\theta^{K_0+2}/(1+\theta)) \int_0^1 ((\theta + (K_0+2)g(x))(1-g(x))/(\theta + g(x))^{K_0+2})\,dx + [p_0 - 1 + (\theta^{K_0+1}/(1+\theta)) \int_0^1 (\theta + (K_0+2)g(x)/(\theta+g(x))^{K_0+1})\,dx] \int_0^1 (g^2(x)(g(x)-1)/(\theta+g(x))^{K_0+3})\,dx/[\int_0^1 (\theta^{K_0+1} g^2(x)/(\theta+g(x))^{K_0+2})\,dx]^{-1}.$$

Use (3.6), (3.7) and (3.8) in (3.11) to simplify it to

(3.12) $$\bar{P} = p_0 + c^2 D\theta^{K_0+2}/(1+\theta)\{(K_0+1)(K_0+2)/(1+\theta)^{K_0+3} + ((2\theta - K_0 - 1)/\theta^{K_0+1}(1+\theta)^2)(p_0 - 1 + \theta^{K_0+1}(\theta + K_0 + 2)/(1+\theta)^{K_0+2})\} + o(c^2).$$

For given θ and p_0, (3.12) gives approximate values of $\bar{P}$ for small c. We abbreviate (3.12) as

(3.13) $$\bar{P} = p_0 + c^2 D d(\theta) + o(c^2).$$

Since for fixed n, the total sample size, the "information" available, and the power must increase with θ, we shall fix $N = n(1+\theta)$ (the total sample size), and find the "best" θ—division of X's and Y's—subject to this constraint. Using (3.13) in (3.5), and writing n as $N/(1+\theta)$ we have for power

(3.14) $1 - \Phi\{[(-Dd(\theta)c^2N^{\frac{1}{2}})/((1+\theta)p_0(1-p_0))^{\frac{1}{2}}] + O(N^{\frac{1}{2}}c^3) + K_\alpha + O(c)\}.$

Clearly, as N increases, this power has a nontrivial limit when

(3.15) $$C = C'/N^{\frac{1}{4}}$$

where C' is arbitrary, and (3.14) then becomes

$$(3.16)\quad 1 - \Phi\{[-Dd(\theta)(C')^2/((1+\theta)p_0(1-p_0))^{\frac{1}{2}}] + K_\alpha + O(N^{-\frac{1}{2}})\}.$$

Limiting power will be maximized when p_0 and θ are chosen to maximize $(d(\theta)/((1+\theta)p_0(1-p_0))^{\frac{1}{2}})$ or to minimize

$$(3.17)\qquad [p_0(1-p_0)(1+\theta)/d^2(\theta)].$$

This minimization is complicated by noting that $d(\theta)$ depends on p_0 through K_0 (see (3.12) and (3.10)).

Table 3.1 shows the behavior of $(1+\theta)/d^2(\theta)$ for $p_0 = \frac{1}{3}$.

TABLE 3.1

θ	K_0	$(1+\theta)/d^2(\theta)$
1	0	103
2	1	69
3	2	65
4	3	68

Since $(1+\theta)/d^2(\theta)$ continues to increase for $\theta > 4$, the optimum θ in this case is 3. This gives $p_0(1-p_0)(65) = 18.9$. Similar computations show that for $p_0 = \frac{1}{2}$, the "best" θ is also 3, giving a $(1+\theta)/d^2(\theta)$ value of 35, with $p_0(1-p_0)$ (35) equalling 8.7. These computations indicate that $\frac{1}{2}$ is a better choice for p_0 than $\frac{1}{3}$, and we conjecture that $p_0 = \frac{1}{2}$ minimizes (3.17).

A reasonable choice of p_0 and θ would then be $p_0 = \frac{1}{2}$ and $\theta = 3$ to give locally good power using this test procedure.

4. Remarks. Analogously to Weiss [6], we can construct a sequential test based on D_n, but the determination of the average sample size properties of that test are quite difficult and the study thereof will be left for a later paper.

I would like to acknowledge the help and inspiration of Professor L. Weiss in the original research on this problem, and the helpful editorial suggestions of Professor I. R. Savage who read an early version. I am also indebted to the referee for his careful and detailed reading and extremely helpful comments.

REFERENCES

[1] Blum, J. R. and Weiss, L. (1957). Consistency of certain two-sample tests. *Ann. Math. Statist.*, **28** 242–246.

[2] Blumenthal, S. (1961). Contributions to the two-sample problem. Technical Report No. 17, Cornell Univ.

[3] Chernoff, H. and Teicher, H. (1958). A central limit theorem for sums of interchangeable random variables. *Ann. Math. Statist.* **29** 118–130.

[4] Hanson, D. L. and Koopmans, L. H. (1965). Convergence rates for the law of large numbers for linear combinations of exchangeable and *-mixing stochastic processes. Technical Report No. 80, Univ. of New Mexico.

[5] Weiss, L. (1955). The stochastic convergence of a function of sample successive differences. *Ann. Math. Statist.* **26** 532–536.

[6] Weiss, L. (1961). Tests of fit based on the number of observations falling in the shortest sample spacings determined by earlier observations. *Ann. Math. Statist.* **32** 838–845.